

# CHAPITRE 3 : L'ANALYSE DEMOGRAPHIQUE DU PASSE

## 3.1 La période de référence

Le modèle théorique, décrit ci-dessus, ne peut toutefois être mis en œuvre que si les diverses séries de quotients intervenant dans les calculs sont déterminées pour chaque zone à projeter. Sur la base de l'axiome, étudier le passé pour prévoir l'avenir, Omphale offre une passerelle entre les données de démographie observées dans le passé (recensements, état civil) et les éléments indispensables pour effectuer un calcul de projection. Mieux encore, Omphale permet de choisir la période du passé qui servira de référence pour ces calculs.

Les comportements migratoires de la population du territoire au cours de la période de référence choisie serviront de base pour le calcul des quotients de migrations nettes. La dernière année (indice a=R2) servira seule de référence pour les quotients de fécondité et de mortalité et sera l'année de départ du calcul de la projection, elle est appelée année de référence. L'année de début de période (indice a=R1) est appelée borne inférieure.

Omphale 2000 offre diverses possibilités de choisir la période de référence :

- Les années des recensements : 1975, 1982, 1990 et 1999 pour tous les zonages.
- Entre 2002 et 20nn, pour les départements et les régions, chaque année pour laquelle on dispose des estimations localisées de population détaillées par sexe et âge.
- À partir de 20nn, chaque année issue du recensement rénové de population. Les modalités de prise en compte de ces résultats ne sont pas encore définies.

## 3.2 La population au premier janvier

Lorsque qu'une borne de la période de référence est un recensement général de la population, il est nécessaire de rétopoler la population de la date du recensement au premier janvier de l'année. Les estimations localisées de populations établies au premier janvier sont directement utilisables par Omphale. Pour le recensement rénové, les modalités d'utilisation ne sont pas encore définies.

Pour les recensements jusqu'en 1999, le département de la démographie de l'Insee élabore la pyramide des âges métropolitaine au premier janvier de l'année du recensement. Le rapport de cette population à la population recensée donne des coefficients de rétopolation par sexe et âge qui sont utilisés par Omphale pour chaque zone. Cette opération préserve la cohérence globale du modèle avec les travaux nationaux.

Si r est l'année du recensement, pour  $0 \leq i \leq 97$

$$P_{z,r,s,i}^* = PRP_{z,r,s,i+1} \cdot P_{.,r,s,i} / PRP_{.,r,s,i+1}$$

La répartition entre les âges 98 et 99 ans au premier janvier implique un calcul distinct car l'effectif de 99 ans au recensement comprend à la fois les personnes qui étaient âgées de 98 ans au premier janvier et les personnes de 99 ans ou plus :

$$P_{z,r,s,98+}^* = PRP_{z,r,s,99} \cdot (P_{.,r,s,98} + P_{.,r,s,99}) / PRP_{.,r,s,99}$$

L'effectif obtenu est réparti entre 98 et 99 ans proportionnellement aux effectifs métropolitains.

Pour  $i=98$  et  $i=99$

$$P_{z,r,s,i}^* = P_{z,r,s,98+}^* \cdot P_{.,r,s,i} / (P_{.,r,s,98} + P_{.,r,s,99})$$

Dans un second temps, la population est ajustée pour tenir compte du mouvement naturel de la zone en supposant un mouvement migratoire nul entre le premier janvier et la date du recensement. Sous ces hypothèses, la population totale de la zone au premier janvier est égale à la population au recensement réduite des enfants nés entre le premier janvier et la date du recensement (classés à 0 an au recensement), et augmentée des décès intervenus au cours de la même période.

Ces décès sont estimés en appliquant aux décès de l'année dans la zone la part des décès avant la date du recensement de l'ensemble de la métropole parmi les décès de l'année ( $\delta$ ). L'état civil pourrait fournir les décès entre le premier janvier et la date du recensement de chaque zone, mais cette solution serait coûteuse en taille de fichier et en temps de calcul pour un gain de qualité limité.

Finalement la population au 1<sup>er</sup> janvier est obtenue par l'équation :

$$P_{z,r,s,i} = P_{z,r,s,i}^* \cdot (PRP_{z,r,s,.} - PRP_{z,r,s,0} + \delta_{.,r} D_{z,r,s,.}) / P_{z,r,s,.}^*$$

**Tableau 2 : Part des décès entre le 1<sup>er</sup> janvier et la date du recensement**

Année	$\delta$
1975	0,18077
1982	0,18077
1990	0,18455
1999	0,20570

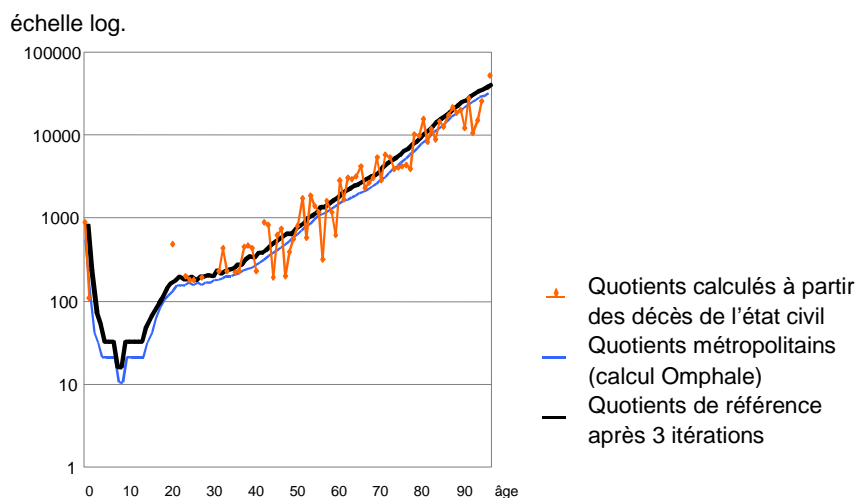
### 3.3 Le calcul de quotients de mortalité de référence

#### 3.3.1 Les bases du calcul

Les quotients de mortalité par sexe et âge sont fortement sensibles aux aléas de la mortalité. Les facteurs conjoncturels, année de forte ou faible mortalité, s'associent à la faiblesse des effectifs de population (aux âges élevés) ou de décès (faible mortalité des jeunes). Si on utilisait les quotients bruts sans précaution, la projection serait très sensible au choix de l'année de référence.

**Figure 1 : Exemple de quotients de mortalité**

Quotients de mortalité d'une zone de 70 000 habitants en 1999  
(quotients pour 100 000 habitants)



Afin de conserver des quotients par âge significatifs, l'application Omphale utilise exclusivement le profil par âge des quotients de mortalité métropolitains. Ces quotients sont transformés par homothétie pour ajuster le nombre de décès par sexe à la mortalité de la zone à projeter.

Pour corriger les aléas conjoncturels, les décès de la zone sont l'objet d'un lissage sur la période de référence. Pour chaque sexe, la part des décès de la zone dans la métropole est régressée par rapport au temps (moindres carrés ordinaires) sur la période de référence. Les décès  $D^*$  utilisés pour les années R1 et R2 ne sont donc pas les décès de l'état civil, mais les décès résultant de la régression.

Pour  $R1 \leq t \leq R2$

$$a = \text{cov}(t, D_{z,t,s..} / D_{.,t,s..}) / \text{var}(t)$$

$$b = \overline{D_{z,t,s..} / D_{.,t,s..}} - a \bar{t}$$

$$D_{z,t,s..}^* = (a \cdot t + b) \cdot D_{.,t,s..}$$

Pour les années  $a=R1$  et  $a=R2$  délimitant la période de référence, on ajuste les quotients de mortalité en multipliant les quotients métropolitains par âge par le rapport des décès lissés aux décès « attendus » : c'est-à-dire le nombre total de décès de chaque sexe qu'il y aurait dans la zone si la mortalité était identique à celle de la métropole (dénominateur de la seconde fraction).

$$QD_{z,a,s,i} = \frac{D_{.,a,s,i}}{P_{.,a,s,i}} \cdot \frac{D_{z,r,s..}^*}{\sum_{i=1}^{99} P_{z,a,s,i} \cdot (1 + QM_{z,a,s,i} / 2) \cdot D_{.,a,s,i} / P_{.,a,s,i}}$$

Dans cette équation, le quotient de mortalité de la métropole est le rapport des décès sur la population sans tenir compte du solde migratoire avec l'étranger. Ce solde est mal connu mais peut être considéré faible par rapport à la population. De plus, aux âges de fortes migrations les quotients de mortalité sont faibles et inversement aux âges de mortalité importante les migrations sont faibles, d'où un effet négligeable sur la courbe des quotients par âge.

Conséquence de ce mode de calcul, le nombre de décès d'une zone pour l'année de référence figurant dans les résultats d'Omphale n'est pas le nombre de décès de l'état civil tel qu'il peut figurer dans d'autres documents : il s'agit du nombre de décès issu de la régression selon un profil par âge identique au profil métropolitain.

Enfin, précisons que le quotient de mortalité à l'âge -1 rapporte les décès d'enfants nés au cours de l'année aux naissances de l'année. Ce n'est pas un taux de mortalité infantile qui rapporte la totalité des décès avant 1 an aux naissances. En effet, une partie des décès avant un an aura lieu l'année suivante.

### 3.3.2 Les quotients de survie

Nous verrons plus loin que le calcul des quotients migratoires nécessite de connaître les quotients de survie (QS) de l'année R1 à l'année R2. Leur calcul fait intervenir les quotients de mortalité de chaque année de la période. Ces quotients de mortalité sont calculés par interpolation linéaire entre R1 et R2 :

$$QD_{z,a,s,i} = QD_{z,R2,s,i} \cdot \frac{a - R1}{R2 - R1} + QD_{z,R1,s,i} \cdot \frac{R2 - a}{R2 - R1}$$

Le quotient de survie de l'année R1 à l'année R2 est le rapport du nombre de survivants d'âge  $i+R2-R1$  au premier janvier de l'année R2 à la population d'âge  $i$  au premier janvier de l'année R1. Le quotient de survie d'une année à la suivante est le complément à un du quotient de mortalité. Pour une période, les quotients annuels se multiplient

entre eux comme des probabilités conditionnelles<sup>1</sup>. Le quotient de survie de R1 à R2 d'une génération d'âge  $i$  l'année R2 est :

$$QS_{z,[R1,R2],s,i} = \prod_{a=R1}^{R2-1} (1 - QD_{z,a,s,i+a-R1})$$

Pour les générations nées au cours des années R1 à R2-1, le quotient de survie ne tient compte que de la période allant de la naissance au premier janvier de l'année R2. Le calcul est similaire en remplaçant R1 par l'année de naissance N et  $i$  par -1 (convention sur l'âge) :

$$QS_{z,[N,R2],s,-1} = \prod_{a=N}^{R2-1} (1 - QD_{z,a,s,-1+a-N})$$

### 3.3.3 Un calcul itératif

Dans l'équation du calcul des quotients de décès, les quotients migratoires (QM) apparaissent alors qu'ils ne sont pas encore déterminés. Inversement, le calcul des quotients migratoires nécessite de connaître les quotients de mortalité.

Omphale procède au calcul des quotients par itération selon le processus suivant :

1. Les quotients de solde migratoire sont initialisés à valeur nulle
2. Les quotients de décès sont calculés en fonction des quotients de migrations nettes
3. Les quotients de migrations nettes sont calculés en fonction des quotients de décès
4. Retour à l'étape 2

Le processus converge rapidement : trois itérations sont suffisantes pour obtenir des quotients stables.

Le calcul des quotients de mortalité n'est pas indépendant de la période de référence choisie ni, à période de référence identique, de la méthode de calcul des quotients de migrations nettes. Lorsque l'on compare des résultats de plusieurs projections d'une même zone basées sur des périodes différentes les écarts sur la mortalité peuvent être sensibles, même pour la somme des décès.

## 3.4 Le calcul des quotients de migrations nettes de référence

### 3.4.1 Les fondements du calcul des quotients

Dans une projection nationale, les migrations sont uniquement les échanges avec l'étranger. Pour une projection locale, ce sont l'ensemble des entrées et sorties du territoire observé qui interviennent, y compris celles, souvent beaucoup plus importantes, avec le reste du territoire national. L'intensité des échanges n'est pas du même ordre de grandeur relativement à la population, aussi les projections locales sont très sensibles aux quotients de migrations nettes.

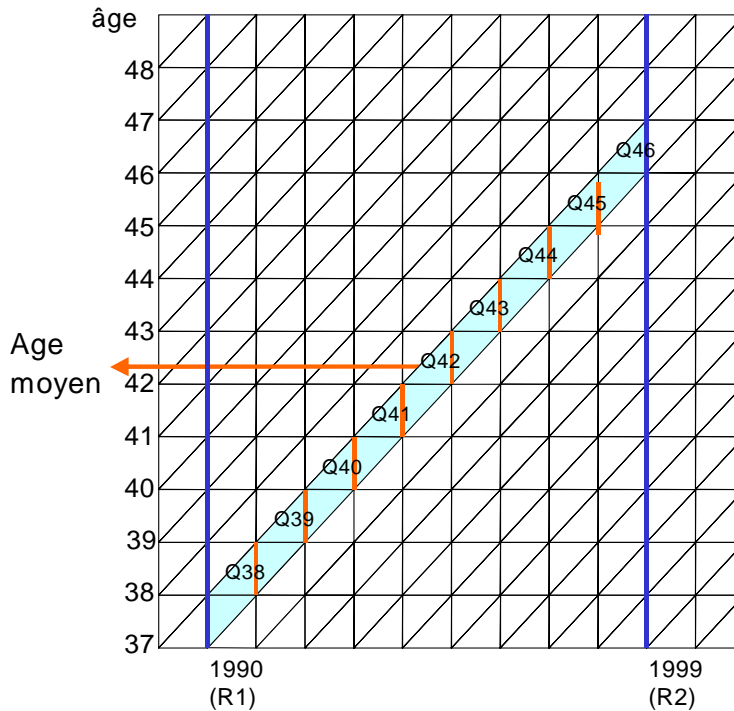
Le calcul des quotients repose sur la comparaison de la population par sexe et âge à la fin de la période de référence avec la population en début de la période de référence réduite de la mortalité. Pour faire des projections par pas annuels, il faut disposer de quotients annuels, alors que les mouvements migratoires ne sont pas connus annuellement. Il est donc nécessaire de faire l'hypothèse simplificatrice que les quotients à sexe et âge donnés

<sup>1</sup> Les quotients sont des fréquences et peuvent donc être interprétés comme des estimations de probabilités.

sont constants tout au long de la période de référence. Par exemple, le quotient des hommes à 45 ans pour la période de référence 1990-1999 est le même pour tous les hommes qui ont atteint cet âge au cours des années 1990 à 1998 : le quotient ne dépend plus de l'année, mais seulement de l'âge et du sexe.

Suivons une génération d'effectif P d'âge i en début de période sur le diagramme de Lexis.

**Figure 2 : Suivi d'une génération sur le diagramme de Lexis**  
Hypothèse de lissage des quotients de migrations nettes



En utilisant l'indice  $a=R$  pour le quotient moyen de la période de référence  $[R1,R2]$ , l'équation 5 (cf. 2.5) donne pour l'année  $R1+1$  un effectif de :

$$P_{z,R1+1,s,i+1} = P_{z,R1,s,i} \cdot (1 - QD_{z,R1,s,i})(1 + QM_{z,R,s,i})$$

pour arriver en fin de période à un effectif de :

$$P_{z,R2,s,i+R2-R1} = P_{z,R1,s,i} \cdot \prod_{a=R1}^{R2-1} (1 - QD_{z,a,s,i+a-R1}) \cdot \prod_{a=R1}^{R2-1} (1 + QM_{z,R,s,i+a-R1})$$

Dans cette équation, le premier produit est le quotient de survie de la période  $[R1,R2]$  calculé précédemment. Pour simplifier davantage l'écriture et par analogie avec la mortalité on définit  $CM=1+QM$ . Finalement, nous obtenons deux équations selon que la génération soit née avant ou après  $R1$ .

**Equation 6 - Lien entre population au début et en fin de période en fonction des quotients :**

$$P_{z,R2,s,i+R2-R1} = P_{z,R1,s,i} \cdot QS_{z,[R1,R2],s,i} \cdot \prod_{a=R1}^{R2-1} CM_{z,R,s,i+a-R1}$$

Pour balayer les générations nées au cours de la période de référence, le calcul ne porte que sur les années allant de l'année de naissance  $N$  à  $R2$ . En utilisant la convention  $i=-1$  pour les enfants nés au cours de l'année et en remplaçant  $R1$  par l'année de naissance  $N$ , nous obtenons une déclinaison de l'équation 6.

### Equation 7 - Lien entre naissances et population en fin de période en fonction des quotients :

$$P_{z,R2,s,-1+R2-N} = P_{z,N,s,-1} \cdot QS_{z,[N,R2],s,-1} \cdot \prod_{a=N}^{R2-1} CM_{z,R,s,-1+a-N}$$

Les quotients de survie, les populations et les naissances étant connus, les équations 6 et 7 donnent un système de 99 équations à 99 inconnues pour chaque sexe. La résolution du système est simple : la seconde équation pour les enfants nés l'année R2-1 (âge 0) donne les quotients de solde migratoire à l'âge -1 (de la naissance au 1<sup>er</sup> janvier de l'année suivante). Ce qui permet d'obtenir les quotients à l'âge 0 pour les enfants nés l'année précédente, et ainsi de suite en remontant d'équation en équation.

Toutefois, ce calcul n'est pas très satisfaisant car les quotients de migrations nettes de tous les âges dépendent directement de la qualité du quotient à l'âge -1. Or ce quotient est fragile, il rapproche les naissances de l'état civil et les effectifs aux recensements, ces deux sources ne présentant pas la même qualité. D'autre part, les approximations sur la saisonnalité des migrations et des naissances jouent le plus sur cet âge et se répercuteraient sur tous les âges. Aussi, la résolution directe du système n'est pas utilisée par Omphale.

### 3.4.2 Le lissage des quotients de migrations nettes

On renonce donc à utiliser les premières équations du système relatives aux générations nées au cours de la période de référence, pour lesquelles on ne peut pas rapprocher les populations en début et fin de période. Le nombre d'inconnues demeurant de 99, la résolution du système d'équations est indéterminée et nécessite des hypothèses supplémentaires.

Une génération, née avant R1, est soumise au cours de la période de référence aux quotients de migrations nettes de (R2-R1) âges successifs. On fait l'hypothèse que tous ces quotients sont représentatifs du quotient à l'âge atteint par la génération au milieu de la période. C'est-à-dire que les n quotients de migrations nettes de la génération au cours de la période sont égaux.

Quand le nombre d'années entre R1 et R2 est impair, l'âge médian  $i$  correspond à la génération qui a l'âge  $i+(R2-R1+1)/2$  en fin de période et  $i-(R2-R1-1)/2$  en début période. Dans l'équation 6, on obtient directement :

$$CM_{z,R,s,i} = \left( \frac{P_{z,R2,s,i+(R2-R1+1)/2}}{P_{z,R1,s,i-(R2-R1-1)/2} \cdot QS_{z,[R1,R2],s,i-(R2-R1-1)/2}} \right)^{\frac{1}{R2-R1}}$$

Quand le nombre d'années entre R1 et R2 est pair, l'âge médian est compris entre deux âges entiers. Dans ce cas, CM à l'âge  $i$  est la moyenne géométrique des deux valeurs qui encadrent l'âge  $i$ .

$$CM1_{z,R,s,i} = \left( \frac{P_{z,R2,s,i+(R2-R1)/2}}{P_{z,R1,s,i-(R2-R1)/2} \cdot QS_{z,[R1,R2],s,i-(R2-R1)/2}} \right)^{\frac{1}{R2-R1}}$$

$$CM2_{z,R,s,i} = \left( \frac{P_{z,R2,s,i+(R2-R1)/2+1}}{P_{z,R1,s,i-(R2-R1)/2+1} \cdot QS_{z,[R1,R2],s,i-(R2-R1)/2+1}} \right)^{\frac{1}{R2-R1}}$$

$$CM_{z,R,s,i} = \sqrt{CM1_{z,R,s,i} \cdot CM2_{z,R,s,i}}$$

Cette méthode est un véritable lissage des quotients de migrations nettes, elle correspond dans l'absolu à une moyenne mobile d'ordre (R2-R1) des quotients réels par âge. Dans la pratique, ce lissage est proche d'une moyenne mobile sur les quotients bruts à la différence près que les quotients bruts sont par hypothèse constants sur la période ce qui n'est pas le cas des quotients réels.

Les quatre méthodes de calcul des quotients de migrations nettes proposées par Omphale utilisent cette méthode de lissage.

### 3.4.3 Les quotients bruts

Pour rétablir l'équilibre entre le nombre d'équations et le nombre de quotients à déterminer, il n'est pas nécessaire de faire l'hypothèse utilisée pour le lissage sur l'ensemble des âges : il suffit de l'appliquer sur une tranche d'âge égale à la durée de la période de référence. De plus, si les quotients de migrations nettes varient peu au sein de la tranche d'âge choisie, l'approximation a un effet réduit.

Omphale propose une tranche d'âge [I1,I2] vers le début de la quarantaine qui présente généralement les caractéristiques souhaitées : effectifs suffisants pour limiter les aléas et stabilité des quotients.

La valeur de I1 peut être sélectionnée dans l'intervalle [ (R2-R1)/2-1 , 98-(R2-R1) ] , par défaut Omphale propose les valeurs :

$$I1 = 42 - (R2 - R1) \div 2$$

$$I2 = I1 + R2 - R1 - 1$$

- Au sein de l'intervalle d'âge I1-I2, les quotients migratoires recherchés sont initialisés par le calcul du lissage présenté précédemment.
- Pour les âges supérieurs à I2, jusqu'à 97 ans, les quotients sont calculés par âge croissant dans l'équation 6 :

$$QM_{z,R,s,i} = \frac{P_{z,R2,s,i+1}}{P_{z,R1,s,i-R2+1+R1} \cdot QS_{z,[R1,R2],s,i-R2+1+R1} \cdot \prod_{a=R1+1}^{R2-1} CM_{z,R,s,i-a+R1}} - 1$$

- Après 97 ans, le calcul des quotients de solde migratoire se complique en raison du regroupement des effectifs de 99 ans ou plus. D'autre part, les effectifs de ces âges sont sujets à de forts aléas. Aussi au-delà de 97 ans les quotients de migrations nettes sont simplement posés égaux à 0.

$$QM_{z,R,s,98} = QM_{z,R,s,99} = 0$$

- Pour les âges inférieurs à I1 et supérieurs à I1+(R2-R1), les quotients sont calculés par âges décroissants dans l'équation 6 :

$$QM_{z,R,s,i} = \frac{P_{z,R2,s,i+R2-R1}}{P_{z,R1,s,i} \cdot QS_{z,[R1,R2],s,i} \cdot \prod_{a=R1+1}^{R2-1} CM_{z,R,s,i+a-R1}} - 1$$

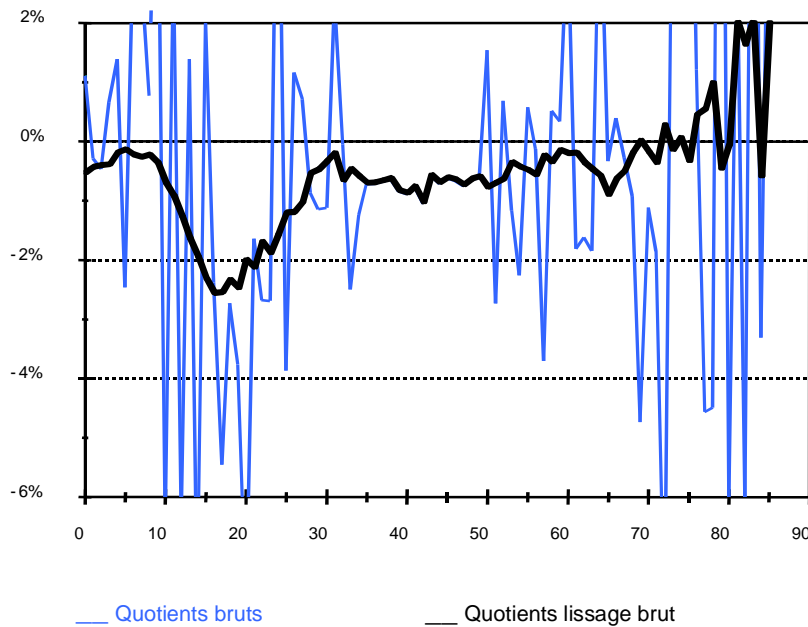
- Pour les quotients des âges inférieurs à R2-R1, l'équation est similaire en utilisant les naissances et quotients de survie de la naissances à R2. Selon la convention sur l'âge -1 et la relation entre l'année de naissance N et l'âge i=N-a-1 :

$$QM_{z,R,s,i} = \frac{P_{z,R2,s,i+R2-N}}{P_{z,N,s,-1} \cdot QS_{z,[N,R2],s,-1} \cdot \prod_{a=N+1}^{R2-1} CM_{z,R,s,-1+a-N}} - 1$$

En théorie, la méthode du calcul brut détermine « exactement » les quotients de solde migratoire, c'est-à-dire que les quotients bruts appliqués à la pyramide des âges au début de la période de référence donne exactement la population en fin de période. Toutefois, en pratique, on n'obtient pas tout à fait ce résultat car les quotients de décès ne sont pas eux-mêmes calculés « exactement ».

Le calcul brut donne des quotients très variables d'un âge à l'autre. Ces fluctuations en dents de scie sont très importantes et présentent des variations périodiques dont le cycle correspond à la durée de la période de référence. Ces considérations nuisent à l'utilisation de ces quotients en projection, ils demeurent néanmoins un bon outil pour étudier le profil par âge des migrations d'un territoire lorsque les méthodes décrites ci-après demandent à être affinées.

**Figure 3 : Exemple de calculs bruts des quotients de migrations nettes**



### 3.4.4 Les quotients lissés bruts

Les quotients lissés bruts sont obtenus par la méthode de lissage décrite en 3.4.3 pour l'ensemble des âges à l'exception des âges extrêmes.

- On annule les quotients pour les âges les plus élevés,

$$QM_{z,R,s,i} = 0$$

- Pour les plus jeunes,  $-1 < i < -1 + (R2 - R1) \div 2$

Le principe de lissage s'applique aux équations 7 en utilisant les naissances au lieu des populations en R1. Ces équations sont en surnombre pour déterminer les quotients manquants. Par analogie avec le traitement des



équations 6, un nombre impair d'années permet le calcul direct des quotients. Aussi seules les générations nées les années  $N=R2-3, R2-5, \dots$  pour lesquelles l'âge  $n$  en milieu de période est entier, sont utilisées.

Compte tenu de la relation entre l'année de naissance  $N$  et l'âge en  $R2$  :  $i=R2-N-1$ , l'équation 7 donne :

$$QM_{z,R,s,i} = \left( \frac{P_{z,R2,s,2i+2}}{P_{z,R2-2i-3,s,-1} \cdot QS_{z,[R2-2i-3,R2],s,-1}} \right)^{\frac{1}{2i+3}} - 1$$

- Pour  $i=-1$

Si la période  $R1-R2$  est d'une année, on utilise la même formule que ci-dessus pour  $i=-1$  :

$$QM_{z,R,s,-1} = \frac{P_{z,R2,s,0}}{P_{z,R2-1,s,-1} \cdot QS_{z,[R2-1,R2],s,-1}} - 1$$

Si la période  $R1-R2$  est de deux années, on corrige la formule pour tenir compte des migrations à l'âge 0 déterminées précédemment (c'est-à-dire le calcul brut) :

$$QM_{z,R,s,-1} = \frac{P_{z,R2,s,1}}{P_{z,R2-2,s,-1} \cdot QS_{z,[R2-2,R2],s,-1} \cdot CM_{z,R,s,0}} - 1$$

Si la période est supérieure à deux ans, la population née l'année précédant le recensement, réputée peu fiable, n'est pas utilisée. Le quotient est obtenu en comparant les populations âgées de 2 à 6 ans en fin de période aux naissances correspondantes.

$$QM_{z,R,s,-1} = \frac{\sum_{k=2}^{\max(6,R2-R1-1)} P_{z,R2,s,k}}{\sum_{k=2}^{\max(6,R2-R1-1)} P_{z,R2-1-k,s,-1} \cdot QS_{z,[R2-1-k,R2],s,-1} \cdot \prod_{i=0}^{k-1} CM_{z,R,s,i}} - 1$$

### 3.4.5 Le lissage médiane moyenne

Les quotients obtenus par le calcul lissé donnent un profil par âge encore un peu accidenté. Pour réduire le bruit, ces quotients « lissés bruts » sont eux-mêmes soumis à une procédure de lissage complémentaire appelée « lissage médiane-moyenne ». Cette procédure effectue une double opération de lissage, une médiane mobile suivie d'une moyenne mobile.

La médiane mobile sur trois années d'âge tend à « raboter » les pointes de la courbe des quotients par âge. Les valeurs excédentaires ne sont pas réparties sur les autres âges. Cependant la médiane mobile conserve, en les inversant, les variations régulières des quotients.

Celles-ci sont aplanies dans un second temps par la moyenne mobile. Il s'agit d'une moyenne géométrique d'ordre trois appliquée à  $CM (=1+QM)$ .

En appelant :

QM les quotients lissés bruts

QMM les quotients lissés par la médiane mobile

QMMM les quotients en sortie de la procédure médiane-moyenne

- Pour  $i=-1$  et  $i=99$  :

$$QMMM_{z,R,s,i} = QMM_{z,R,s,i} = QM_{z,R,s,i}$$

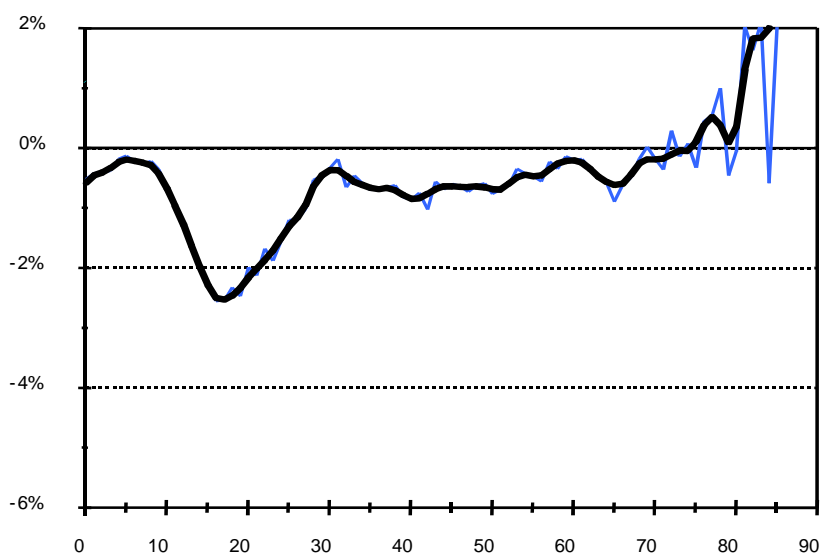
- Pour  $0 \leq i \leq 98$  :

$$QMM_{z,R,s,i} = \text{médiane}(QM_{z,R,s,i-1}, QM_{z,R,s,i}, QM_{z,R,s,i+1})$$

$$QMMM_{z,R,s,i} = (1 - QMM_{z,R,s,i-1})(1 - QMM_{z,R,s,i})^2(1 - QMM_{z,R,s,i+1})^{1/4} - 1$$

Le lissage médiane-moyenne donne généralement un profil migratoire par âge qui se prête aisément à l'interprétation. Le profil donné en exemple est caractéristique d'une zone qu'une partie des jeunes quittent pour leurs études ou pour trouver un premier emploi.

**Figure 4 : Exemple de quotients de migrations nettes après lissage médiane-moyenne**



— Quotients lissage brut    — Quotients lissage médiane-moyenne

Cependant cette méthode a un inconvénient : les lissages successifs tendent par construction à étaler sur les âges voisins des comportements migratoires caractéristiques d'une tranche d'âge bien précise. Ainsi dans l'exemple, le creux dû à l'émigration des jeunes commence en données lissées dès 10 ans alors que les départs réels se produiraient plutôt vers l'entrée en seconde ou plus probablement encore après le baccalauréat. Le même phénomène se produit dans l'autre sens quelque part après 20 ans.

### 3.4.6 Le lissage-délissage

Pour corriger les défauts du lissage médiane-moyenne, une technique de « délissage » de ces quotients a été développée dans Omphale. Dans les anciennes versions, cette technique pouvait s'utiliser sous deux formes : le délissage ascendant et le délissage descendant. Dans la nouvelle version de l'application, toutes ces techniques sont intégrées dans un seul module « de lissage composite » que nous verrons après avoir exploré la méthode du « lissage-délissage ».

Nous avons vu dans le traitement du lissage (3.4.2) que le quotient lissé pour un âge  $i$  est la moyenne des quotients réels d'une tranche d'âge  $i+n$  à  $i-n$ . En négligeant les variations temporelles des quotients par âge, la

différence entre le quotient lissé de l'âge  $i$  et celui de  $i+1$  représente la différence entre les quotients des âges extrêmes  $i-n$  et  $i+n+1$ . Si l'un de ces deux quotients est connu, l'autre s'en déduit.

Si l'on considère que les quotients obtenus par le lissage médiane-moyenne sont les « bons » quotients pour les enfants jusqu'à un âge  $A$ , alors on peut calculer un nouveau quotient pour l'âge  $A+1$ . Ce nouveau quotient permet à son tour de calculer un nouveau quotient à l'âge  $A+2$ , et ainsi de suite pour les âges suivants : c'est le calcul proposé par le délissage ascendant.

De manière symétrique, si l'on considère que le lissage médiane-moyenne donne les « bons » quotients à partir de l'âge  $C$ , on peut calculer un nouveau quotient pour l'âge  $C-1$ . Cette méthode opérant par âge décroissant est appelée délissage descendant.

Le développement du calcul des quotients déliés est plus simple en utilisant CMMM (=1+QMMM). Comme dans le calcul du lissage, le nombre pair ou impair d'années de la période de référence nécessite un ajustement du calcul :

### a - Cas où $R2-R1$ est impair :

Le quotient à l'âge  $i+(R2-R1-1)/2$  est la moyenne des quotients des âges  $i$  à  $(i-1+R2-R1)$  soit :

à l'âge  $i$  :

$$CMMM_{z,R,s,i+(R2-R1-1)/2} = \left( \prod_{k=i}^{i-1+R2-R1} CMMM_{z,R,s,k} \right)^{\frac{1}{R2-R1}}$$

à l'âge  $i+1$  :

$$CMMM_{z,R,s,i+(R2-R1+1)/2} = \left( \prod_{k=i+1}^{i+R2-R1} CMMM_{z,R,s,k} \right)^{\frac{1}{R2-R1}}$$

le rapport de la deuxième équation sur la première donne après simplification :

$$\left( \frac{CMMM_{z,R,s,i+(R2-R1+1)/2}}{CMMM_{z,R,s,i+(R2-R1-1)/2}} \right)^{R2-R1} = \left( \frac{CMMM_{z,R,s,i+R2-R1}}{CMMM_{z,R,s,i}} \right)$$

C'est à dire pour le délissage descendant :

$$CMMM_{z,R,s,i} = CMMM_{z,R,s,i+R2-R1} \left( \frac{CMMM_{z,R,s,i+(R2-R1-1)/2}}{CMMM_{z,R,s,i+(R2-R1+1)/2}} \right)^{R2-R1}$$

et pour le délissage ascendant en posant  $j=i+R2-R1$  :

$$CMMM_{z,R,s,j} = CMMM_{z,R,s,j-R2+R1} \left( \frac{CMMM_{z,R,s,j-(R2-R1-1)/2}}{CMMM_{z,R,s,j-(R2-R1+1)/2}} \right)^{R2-R1}$$

## b - Cas où R2-R1 est pair :

L'âge en milieu de période n'est pas un âge entier, comme dans le cas du lissage on procède au calcul sur les deux âges qui encadrent l'âge cherché puis on prend la moyenne géométrique des CM ce qui donne :

Pour le lissage descendant :

$$CMMM_{z,R,s,i} = CMMM_{z,R,s,i+R2-R1} \left( \frac{CMMM_{z,R,s,i+(R2-R1-2)/2}}{CMMM_{z,R,s,i+(R2-R1+2)/2}} \right)^{\frac{R2-R1}{2}}$$

et pour le délissage ascendant :

$$CMMM_{z,R,s,j} = CMMM_{z,R,s,j-R2+R1} \left( \frac{CMMM_{z,R,s,j-(R2-R1-2)/2}}{CMMM_{z,R,s,j-(R2-R1+2)/2}} \right)^{\frac{R2-R1}{2}}$$

### 3.4.7 Le lissage composite

Le lissage composite est la forme la plus aboutie du calcul des quotients de migrations nettes dans le développement de l'application Omphale. Son principe est d'appliquer à chaque tranche d'âge la méthode de calcul la plus appropriée aux caractéristiques des migrations pour ces âges.

L'utilisateur peut définir 4 bornes (A, B, C et D) qui délimitent les 5 tranches d'âge au maximum. Chaque tranche d'âge sera traitée par une méthode de calcul particulière :

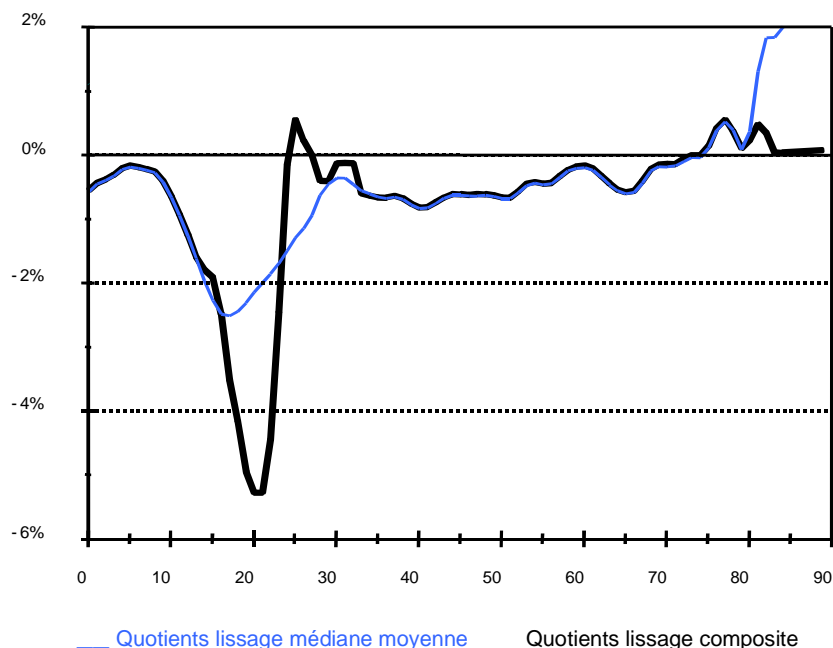
- Pour les enfants  $-1 \leq i \leq A-3$  Lissage médiane moyenne
- Pour les adolescents  $A+3 \leq i \leq B-3$  Lissage-délissage ascendant
- Pour les jeunes adultes  $B+3 \leq i \leq C-3$  Lissage-délissage descendant
- Pour les adultes  $C+3 \leq i \leq D-3$  Lissage médiane moyenne
- Pour les plus âgés  $D+3 \leq i \leq 99$  Quotient nul (aléas)

Les valeurs de A, B, C et D sont soumises à des contraintes :

- $A \geq R2-R1+1$  : la méthode de délissage ascendant nécessite de « valider » les quotients issus du lissage médiane moyenne jusqu'au premier âge « délissable » correspondant à l'âge -1, c'est-à-dire  $R2-R1+1$ .
- $B \geq A+6$  : pour permettre le calcul de l'interpolation
- $97-R2+R1 \geq C \geq B+6$  pour permettre le calcul des interpolations et l'application du délissage descendant
- $96 \geq D \geq C+6$  pour permettre les interpolations

L'analyse des quotients bruts peut apporter une aide à la détermination des bornes les plus adaptées au zonage. Par défaut, Omphale propose 12,18,24 et 85 ans si la période de référence le permet.

**Figure 5 : Comparaison des quotients de migrations nettes selon les méthodes de lissage**



Lorsque la période de référence est longue les contraintes sur les valeurs de A et B peuvent empêcher d'appliquer correctement le lissage composite.

C'est le cas dans l'exemple donné en figure 5 où la période de référence est R1=1975 et R2=1990. Les quotients obtenus par le lissage médiane-moyenne semblent trop élevés dès l'âge de 10 ans, mais le délissage ascendant ne peut pas être effectué avant l'âge de 16 ans (valeur minimale permise pour A).

Ce problème se répercute d'ailleurs sur les bornes suivantes. Les contraintes sur les valeurs de B et C ne permettent que des valeurs trop élevées, ce qui génère des fluctuations de la courbe des quotients entre 25 et 35 ans.

L'application permet de ne pas renseigner certaines bornes ce qui permet d'appliquer chacune des méthodes à d'autres populations. La méthode de calcul pour chaque âge selon les diverses possibilités est donnée dans le tableau 3.

Pour éviter les décrochages lors du changement de méthode de calcul, autour de l'âge où intervient le changement les quotients sont calculés par interpolation entre les quotients obtenus selon les deux méthodes.

Pour un passage à l'âge X d'une méthode donnant un quotient QM1 à une méthode donnant un quotient QM2, les quotients pour l'intervalle [X-2, X+2] sont calculés par :

$$CMI_{z,R,s,X+k} = \left( CMI_{z,R,s,X+k}^{3-k} \cdot CMI_{z,R,s,X+k}^{3+k} \right)^{\frac{1}{6}}$$

Dans la pratique, l'interpolation laisse encore quelques décrochages pour les raccordements des données délissées, une médiane mobile d'ordre trois est finalement appliquée sur les quotients des tranches d'âges [A-3,B-3] et [B+3,C+3].

**Tableau 3 : Méthode de calcul des quotients de migrations nettes selon l'âge pour le lissage composite**

Tranche d'âge	Bornes			
	A,B,C,D	_,B,C,D	A,_,C,D	_,_,_,D
-1<=i<=A-3	QMMM	QMMM	QMMM	QMMM
A-2<=i<=A+2	interpolation	QMMM	interpolation	QMMM
A+3<=i<=B-3	QMDA	QMMM	QMDA	QMMM
B-2<=i<=B+2	interpolation	interpolation	QMDA	QMMM
B+3<=i<=C-3	QMDD	QMDD	QMDA	QMMM
C-2<=i<=C+2	interpolation	interpolation	interpolation	QMMM
C+3<=i<=D-3	QMMM	QMMM	QMMM	QMMM
D-2<=i<D+2	interpolation	interpolation	interpolation	interpolation
D+3<=i<=99	0	0	0	0

QMMM= médiane moyenne  
 QMDA= délissage ascendant  
 QMDD= délissage descendant  
 En encadré, les tranches d'âge lissées par médiane mobile d'ordre 3 en fin de calcul

### 3.4.8 Les redressements manuels et autres possibilités

Les méthodes de calculs des quotients de migration nette sont construites pour répondre à une demande pour toute zone du territoire. En raison de spécificités locales portant sur les données (naissances nombreuses à l'étranger qui ne font pas l'objet d'une domiciliation par exemple) ou sur les comportements migratoires, ces méthodes peuvent ne pas donner de résultats satisfaisants pour des zones particulières.

Omphale offre une solution à ces problèmes par une correction « manuelle » des quotients migratoires. Cette solution est plutôt à réserver à des corrections ponctuelles, car elle est assez difficile à mettre en œuvre. Une autre méthode, qui peut répondre à des besoins d'analyses, est l'utilisation en projection de quotients de référence calculés sur un autre territoire que le territoire à projeter. C'est-à-dire :

$$QM_{z,R,s,i} = QM_{z',R',s,i}$$

## 3.5 Le calcul des quotients de fécondité de référence

### 3.5.1 Les bases du calcul

Comme pour la mortalité, les quotients de fécondité par âge des mères sont sensibles aux aléas conjoncturels ou à la faiblesse des effectifs. Omphale remédie à ce problème par une méthode analogue aux décès.

Afin de conserver des quotients par âge significatifs, l'application Omphale utilise exclusivement le profil par âge des quotients de fécondité métropolitains. Ces quotients sont transformés par homothétie pour ajuster le nombre de naissances à la natalité de la zone à projeter.

Pour corriger les aléas conjoncturels, les naissances de la zone sont calculées par lissage sur la période de référence. La part des naissances de la zone dans la métropole est régressée (moindres carrés ordinaires) sur la période de référence pour obtenir des naissances lissées ( $N^*$ ) l'année R2.

Pour  $R1 \leq t \leq R2$

$$a = \text{cov}(t, N_{z,t,\dots} / N_{.,t,\dots}) / \text{var}(t)$$

$$b = \overline{N_{z,t,\dots} / N_{.,t,\dots}} - a \bar{t}$$

$$N_{z,t,\dots}^* = (a \cdot t + b) \cdot N_{.,t,\dots}$$

Comme pour les décès, le nombre de naissances de l'année de référence figurant dans les résultats d'Omphale n'est donc pas le nombre de naissances de l'état civil.

Seuls les quotients de fécondité de l'année de référence sont nécessaires. Ils sont obtenus en multipliant les quotients de fécondité par âge métropolitains par le rapport des naissances lissées aux naissances « attendues » : c'est-à-dire le nombre total de naissances qu'il y aurait dans la zone si la fécondité était identique à celle de la métropole.

Pour les naissances, il ne faut pas oublier que l'indice  $s$  concerne le sexe de l'enfant et l'âge  $j$  est l'âge de la mère (dénominateur de la seconde fraction).

$$QF_{z,R2,j} = \frac{N_{.,R2,j}}{P_{.,R2,j}} \cdot \frac{N_{z,R2,\dots}^*}{\sum_{i=15}^{49} P_{z,R2,i} \cdot (1+QM_{z,R2,i}/2) \cdot (1-QD_{z,R2,i}/2) \cdot \frac{N_{.,R2,i}}{P_{.,R2,i}}}$$

Les quotients migratoires pour l'année R2 sont supposés égaux aux quotients annuels moyens de la période de référence. Seuls les quotients de fécondité de l'année R2 sont nécessaires au fonctionnement d'Omphale.

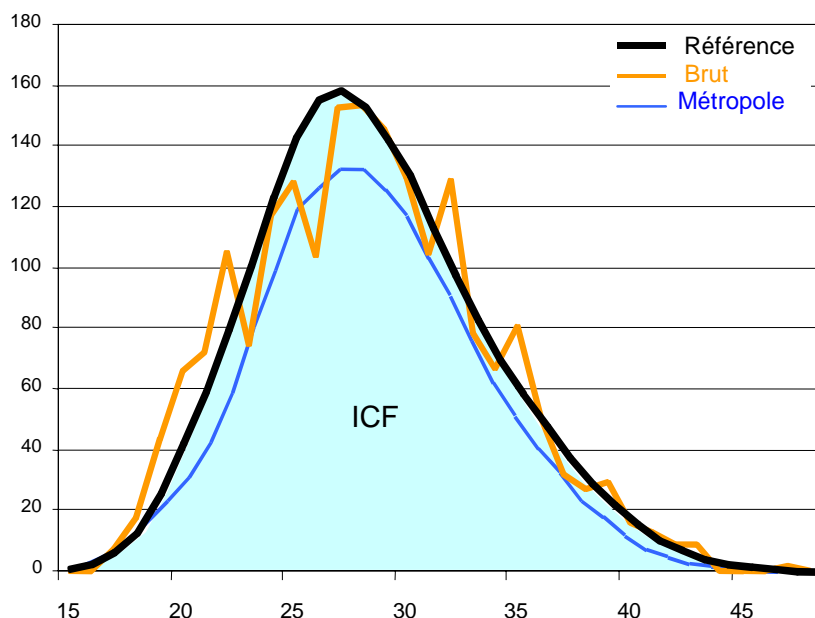
### 3.5.2 L'indicateur conjoncturel de fécondité

Les quotients de fécondité, à la différence des quotients de mortalité, peuvent être additionnés. La somme des quotients sur tous les âges est la somme des naissances réduites plus généralement appelée indicateur conjoncturel de fécondité (abrégié ICF) : c'est un nombre fictif d'enfants que mettrait au monde une femme tout aussi fictive qui toute sa vie féconde connaîtrait les conditions de fécondité de l'année considérée. Sur le graphique c'est la surface comprise entre l'axe horizontal et la courbe des quotients :

$$ICF_{z,a} = \sum_{j=15}^{49} QF_{z,a,2,j}$$

Cet indicateur sera très pratique pour formuler des hypothèses d'évolution de la fécondité. L'inconvénient est qu'il joue uniquement sur la surface sous la courbe mais non sur sa forme.

Figure 6 : Exemples de courbes de taux de fécondité par âge



### 3.6 Le calcul des espérances de vie

Un indicateur souvent utilisé pour caractériser la mortalité est l'espérance de vie ou durée moyenne de vie. Omphale n'utilise pas directement l'espérance de vie pour projeter les populations, le calcul est effectué *a posteriori*, en utilisant les décès et les populations projetés.

L'espérance de vie, une année donnée, est la durée moyenne de vie d'une génération fictive soumise à chaque âge aux quotients de mortalité de cette année. Le calcul présenté ici n'est pas aussi précis que le calcul classique tenant compte des décès avant et après anniversaire, mais il en donne une bonne approximation.

Soit une génération fictive de  $K$  individus nés l'année  $A$ . Le nombre de décès en cours d'année s'obtient en appliquant le quotient de mortalité ( $QD=1-QS$ ) aux survivants en début d'année :

$$D_{z,A+n,s,n-1} = (1 - QS_{z,A,s,n-1}) \cdot K \cdot \prod_{i=-1}^{n-2} QS_{z,A,s,i}$$



Chacune des personnes décédées au cours de l'année A+n aura vécu n années :

- 0,5 année de la naissance au 1/1/A+1 (équirépartition des naissances sur l'année)
- 0,5 année du 1/1/A+n au décès (équirépartition des décès sur l'année)
- n-1 années entre le 1/1/A+1 et le 1/1/A+n.

Soit, pour l'ensemble des personnes décédées au cours de l'année A+n un nombre d'années vécues de :

$$X_{z,A+n,s,n-1} = n \cdot K \cdot (1 - QS_{z,A,s,n-1}) \cdot \prod_{i=-1}^{n-2} QS_{z,A,s,i}$$

La durée moyenne de vie est obtenue en sommant les valeurs de X pour tous les âges et en divisant par K. Pour les décès de la naissance au premier janvier A+1 (i=-1), on retient par convention une durée moyenne de vie de 0,25 année (cf. Annexe 3).

En supposant que la génération soit totalement éteinte l'année A+n<sub>max</sub>, le nombre moyen d'années vécues par la génération est :

$$E_{z,A,s} = 0,25 - 0,25 \cdot QS_{z,A,s,-1} + \sum_{n=1}^{n_{\max}} n \cdot \left( \prod_{i=-1}^{n-2} QS_{z,A,s,i} - \prod_{i=-1}^{n-1} QS_{z,A,s,i} \right)$$

En développant la somme, n s'élimine par soustraction d'un âge au suivant et comme le quotient de survie de la dernière année est nul (sinon ce n'est pas la dernière année !), l'équation devient :

$$E_{z,A,s} = 0,25 + 0,75 \cdot QS_{z,A,s,-1} + \sum_{n=1}^{n_{\max}} \prod_{i=-1}^{n-1} QS_{z,A,s,i}$$

Dans la pratique, Omphale ne dispose que des quotients de survie jusqu'à 99 ans, or les effectifs de centenaires sont de plus en plus nombreux et leur poids dans le calcul de l'espérance de vie est de moins en moins négligeable<sup>2</sup>. Pour palier le manque d'information, on reconduit le quotient de survie à 99 ans (de l'ordre de 0,7 %) aux âges suivants jusqu'à 115 ans. Ce qui donne finalement :

$$E_{z,A,s} = 0,25 + 0,75 \cdot QS_{z,A,s,-1} + \sum_{n=1}^{99} \prod_{i=-1}^{n-1} QS_{z,A,s,i} + \sum_{n=1}^{15} QS_{z,A,s,99}^n \cdot \prod_{i=-1}^{98} QS_{z,A,s,i}$$

Ce calcul très lourd de l'espérance de vie est différent de l'approximation de l'ancienne version d'Omphale. Il est plus conforme au calcul classique de l'espérance de vie que l'on peut trouver dans d'autres ouvrages comme « La situation démographique en..... » par exemple.

<sup>2</sup> L'effet de levier sur la moyenne des personnes loin de la valeur centrale accentue le poids des centenaires dans le calcul de l'espérance de vie.